Description of a
Three Day Short Course

Techniques of Geostatistical Estimation and Simulation Applied to Environmental Geology

Prepared for
U.S. Nuclear Regulatory Commission
December, 1997

Paul Kaplan
Sandia National Laboratories
PO Box 5800, MS-0716
Albuquerque, NM 87185
505-284-4786

Smart Sampling

Background

The technical short course titled "Techniques of Geostatistical Estimation and Simulation Applied to Environmental Geology" was prepared for and given at the Geological Society of America Annual Meeting in October 1997. The course integrates exploratory data analysis, advanced geostatistical simulation, and economic-risk assessment. The integration of these disciplines is based on lessons learned as a project called SmartSampling that was sponsored by the U.S. Department of Energy's Office of Science and Technology and by the Innovative Treatment Remediation Demonstration Program, a public-private partnership developed in cooperation with the U.S. Environmental Protection Agency's Technology Innovation Office.

Course Description

Environmental scientists are often involved in making economic, remedial, and regulatory decisions regarding the amount and distribution of spatially heterogeneous contaminants based upon limited sample data. This course will introduce the student to a suite of geostatistical tools and economic models that allow environmental scientists to model and use the spatial correlation inherent in most geotechnical data sets to build maps, quantify the uncertainty in those maps, and evaluate the economic and regulatory consequences of alternative remedial design strategies. The techniques covered in this course include exploratory data analysis, calculation and modeling of spatial correlation with variograms, spatial estimation through kriging, spatial simulation techniques, uncertainty analysis, and economic analysis. Example problems from environmental sites will be evaluated with hands on instruction using public domain software. The course stresses geostatistics, economic-risk analysis, probabilistic analyses, and environmental decision making.

Course Outline

The course is divided into nine technical modules including a case study. After introductory and background material, the modules are presented in the following order:

1. **Case Study, Mound PRS-379**
2. **Exploratory Data Analysis**
    - Mapping the data set
    - Histogram techniques
    - Probability-plotting
    - Correlation between multivariate data
    - Data transformations (logarithmic, indicator, rank-order, normal score, uniform-score, trend analysis and removal)
3. **Spatial Correlation**
    - Calculation of experimental variograms
    - Fitting models to variograms
    - Bi-gaussian check

- Concepts of anisotropy and nested structures in variograms
- Other techniques for defining spatial variability

4. **Spatial Estimation**
   - Review of techniques (nearest neighbor, inverse distance, trend surface, splines)
   - Kriging, indicator kriging, probability kriging
   - Checking kriging output
5. **Simulation**
   - Difference between estimation and simulation
   - Basics of probabilistic risk assessment
   - Transfer of uncertainty in spatial distribution to uncertainty in decision-making
   - Testing the model
6. **Probability Mapping**
   - Concept of the probability of exceedence
   - Concept of probability mapping
   - Incorporation of spatial uncertainty in remediation maps
   - Cross validation
7. **Economics**
   - Cost elements and assumptions
   - Objective functions
   - Linear loss function
   - Squared loss function
   - Data worth
   - Ranking and location of potential additional sample sites
   - Probabilistic calculation of regulatory action levels
   - Economic risk
8. **Scaling**
   - Measurement versus remediation scales
   - Analytic techniques
   - Numerical techniques
9. **Summary**

Problem sets to accompany the course material are provided along with a manual and a disk containing public domain software that will be used during the class. The course stresses geostatistics, economic-risk analysis, probabilistic analyses, and environmental decision making. Prior to the course, a questionnaire will be sent to each of the participants. This questionnaire will be returned to the instructors a month or two before the scheduled date of the course. The answers will be used to tailor specific sections of the course to the skill levels of the students.

Instructors

Three instructors will be provided to teach the modules and to work closely with the students on the problem sets.

Sean A. McKenna, Sandia National Laboratories, Albuquerque, New Mexico. Sean is a graduate of Colorado School of Mines and one of the co-authors of UNCERT, a public domain software package for exploratory data analysis and visualization. Sean is currently involved in modeling stochastic rock properties for the Yucca Mountain Project, developing techniques for constraining stochastic simulations of hydrologic properties to well test and tracer transport data for the WIPP project, and creation of remediation maps and second-phase sample optimization using geostatistics for contaminated soil problems.

Christopher A. Rautman, Sandia National Laboratories, Albuquerque, New Mexico. Chris is a graduate of the University of Wisconsin in Madison. Chris is currently the task leader for three-dimensional rock property modeling at the Yucca Mountain site. He is the author of numerous papers on the application of advanced geostatistical techniques to probabilistic decision-making at environmental contamination sites.

Paul G. Kaplan, Sandia National Laboratories, Albuquerque, New Mexico. Paul is a graduate of Purdue University. He is currently manager of the SmartSampling project.